



G

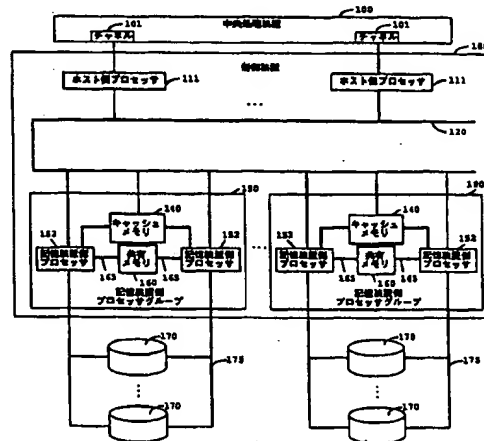
| | | |
|---|-----------|--|
| <p>(51) 国際特許分類6 G06F 3/06</p> | <p>A1</p> | <p>(11) 国際公開番号 WO98/59291</p> <p>(43) 国際公開日 1998年12月30日 (30.12.98)</p> |
| <p>(21) 国際出願番号 PCT/JP97/02135</p> <p>(22) 国際出願日 1997年6月20日 (20.06.97)</p> <p>(71) 出願人 (米国を除くすべての指定国について) 株式会社 日立製作所(HITACHI, LTD.)(JP/JP) 〒101 東京都千代田区神田駿河台四丁目6番地 Tokyo, (JP)</p> <p>(72) 発明者; および (75) 発明者/出願人 (米国についてのみ) 山神憲司(YAMAGAMI, Kenji)(JP/JP) 〒228 神奈川県相模原市豊町17-12 日立上鶴間社宅A501 Kanagawa, (JP)</p> <p>山本 彰(YAMAMOTO, Akira)(JP/JP) 〒229 神奈川県相模原市若松6丁目5番61号 Kanagawa, (JP)</p> <p>(74) 代理人 弁理士 小川勝男(OGAWA, Katsuo) 〒100 東京都千代田区丸の内一丁目5番1号 株式会社 日立製作所内 Tokyo, (JP)</p> | | <p>(81) 指定国 CN, JP, KR, US, 欧州特許 (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>添付公開書類 国際調査報告書</p> |

(54) Title: METHOD OF CONTROL OF MEMORY CONTROLLER

(54) 発明の名称 記憶制御装置の制御方法

(57) Abstract

One or more memory-device-side processors which can access the memory device belong to one group. Common memories and cache memories are allotted to the respective groups. When the access to the memory device is requested by a host-side processor, the memory-device-side processor of the group to which the memory device belongs informs the host-side processor of the stored cache address or the cache address to be stored and the data length. The host-side processor reads data at the address or writes data into the address. Thus, the access neck of the common memory and the cache memory is reduced and the system throughput is improved.



- 100 ... Central processing unit
- 101 ... channel
- 111 ... host-side processor
- 140 ... cache memory
- 182 ... memory-device-side processor
- 160 ... common memory
- 180 ... controller
- 190 ... memory-device-side processor group

BEST AVAILABLE COPY

(57)要約

記憶装置にアクセス可能な1つ以上の記憶装置側プロセッサを同一のグループとして、各グループに共有メモリおよびキャッシュメモリを割り当てる。ホスト側プロセッサから記憶装置へのアクセス要求に対して、この記憶装置が属するグループの記憶装置側プロセッサは、データを格納しているか、あるいは格納すべきキャッシュアドレスおよびデータ長をホスト側プロセッサに通知する。そして、ホスト側プロセッサはこのアドレスからのデータ読み出し、あるいは当該アドレスへデータ書き込みを実行する。これにより、共有メモリおよびキャッシュメモリのアクセスネックを低減し、システムスループットを向上させることができる。

PCTに基づいて公開される国際出願のパンフレット第一頁に掲載されたPCT加盟国を同定するために使用されるコード(参考情報)

| | | | | | | | |
|----|--------------|----|-----------|----|----------------|----|------------|
| AL | アルバニア | FI | フィンランド | LK | スリ・ランカ | SI | スロヴェニア |
| AM | アルメニア | FR | フランス | LR | リベリア | SK | スロヴァキア |
| AT | オーストリア | GA | ガボン | LS | レソト | SL | シエラ・レオネ |
| AU | オーストラリア | GB | 英国 | LT | リトアニア | SN | セネガル |
| AZ | アゼルバイジャン | GD | グレナダ | LU | ルクセンブルグ | SZ | スワジランド |
| BA | ボスニア・ヘルツェゴビナ | GE | グルジア | LV | ラトヴィア | TD | チャード |
| BB | バルバドス | GH | ガーナ | MC | モナコ | TC | トーゴ |
| BE | ベルギー | CM | ガンビア | MD | モルドヴァ | TJ | タジキスタン |
| BF | ブルキナ・ファソ | CN | ギニア | MG | マダガスカル | TM | トルクメニスタン |
| BG | ブルガリア | CW | ギニア・ビサウ | MK | マケドニア旧ユーゴスラヴィア | TR | トルコ |
| BJ | ベナン | GR | ギリシャ | | 共和国 | TT | トリニダード・トバゴ |
| BR | ブラジル | HR | クロアチア | ML | マリ | UA | ウクライナ |
| BY | ベラルーシ | HU | ハンガリー | MN | モンゴル | UG | ウガンダ |
| CA | カナダ | ID | インドネシア | MR | モーリタニア | US | 米国 |
| CF | 中央アフリカ | IE | アイルランド | MW | マラウイ | UZ | ウズベキスタン |
| CG | コンゴ | IL | イスラエル | MX | メキシコ | VN | ヴェトナム |
| CH | スイス | IN | インド | NE | ニジェール | YU | ユーゴスラビア |
| CI | コートジボワール | IS | アイスランド | NL | オランダ | ZW | ジンバブエ |
| CM | カメルーン | IT | イタリア | NO | ノルウェー | | |
| CN | 中国 | JP | 日本 | NZ | ニュージーランド | | |
| CU | キューバ | KE | ケニア | PL | ポーランド | | |
| CY | キプロス | KG | キルギスタン | PT | ポルトガル | | |
| CZ | チェコ | KP | 北朝鮮 | RO | ルーマニア | | |
| DE | ドイツ | KR | 韓国 | RU | ロシア | | |
| DK | デンマーク | KZ | カザフスタン | SD | スーダン | | |
| EE | エストニア | LC | セントルシア | SE | スウェーデン | | |
| ES | スペイン | LI | リヒテンシュタイン | SG | シンガポール | | |

明 細 書

記憶制御装置の制御方法

5 技術分野

本発明は、中央処理装置と記憶装置との間のデータ転送に介在する制御装置の制御方法に関する。

背景技術

- 10 従来の制御装置は、記憶装置のデータを一時的に格納するキャッシュメモリと、キャッシュメモリと中央処理装置との間のデータ転送を実行するホスト側プロセッサと、キャッシュメモリと記憶装置との間のデータ転送を実行する記憶装置側プロセッサと、キャッシュメモリの管理情報を格納する共有メモリから構成される。そして、ホスト側プロセッサ
15 と記憶装置側プロセッサの双方が、共有メモリにアクセス可能となっている。

- 上記の構成では、制御装置内の各プロセッサは、キャッシュメモリ上のデータおよび管理情報を共有している。このため、各プロセッサは、キャッシュメモリを介していずれの記憶装置のデータにもアクセス可能
20 である。また、ホスト側及び記憶装置側の各プロセッサは、独立して動作可能である。つまり、ホスト側プロセッサによるキャッシュメモリへのヒット処理と、記憶装置側プロセッサによる記憶装置へのライト処理は、並列に実行可能である。

- 従来のように、制御装置に一つの共有メモリ及びキャッシュメモリを
25 搭載する構成では、各プロセッサからのアクセスが共有メモリに集中する。このため、共有メモリへのアクセス制御がボトルネックとなるとい

う問題がある。

本発明の目的は、共有メモリおよびキャッシュメモリへのアクセスネックを解消する記憶制御装置およびその制御方法を提供することにある。

5

発明の開示

キャッシュメモリは記憶装置上のデータを一時的に保存しておくためのものであり、基本的には記憶装置対応に存在すれば良い。また、共有メモリに格納されたキャッシュ管理情報も同様に、記憶装置対応に存在すれば良い。以上の考えに基づいて、上記目的を達成するために本発明では、各記憶装置毎に、この記憶装置にアクセス可能な1つ以上の記憶装置側プロセッサを一つのグループとして割り当て、キャッシュメモリおよび共有メモリをそれぞれのグループに割り当て、グループ内の各記憶装置側プロセッサが、このグループに割り当てられたキャッシュ管理情報10 情報を管理する。これにより、複数のキャッシュメモリおよび複数の共有メモリを同時に動作させることが可能となり、ボトルネックを解消することができる。ただし、このグループ分けは、複数の記憶装置に対するグループ分けであっても良い。

本発明における制御装置では、ホスト側プロセッサは、中央処理装置20 からアクセス要求のあったデータが格納された記憶装置に基づいて、その記憶装置にアクセス可能なグループ内の記憶装置側プロセッサを特定する。また、特定された記憶装置側プロセッサは、ホスト側プロセッサからのアクセス要求に対して、このグループに割り当てられた共有メモリにアクセスして、目的とするデータがキャッシュ上に存在するかどうかを判定し、キャッシュにデータが存在すれば、データを格納している25 キャッシュアドレスをホスト側プロセッサに通知する。この結果、ホス

ト側プロセッサは、通知されたキャッシュアドレスに対してアクセスする。

以上の処理により、各グループに割り当てられた1つの共有メモリへアクセスするのは、この共有メモリが割り当てられた1つのグループに
5 属する記憶装置側プロセッサのみである。このため、1つの共有メモリに対するアクセス回数を削減できる。この結果、共有メモリに対するアクセスの競合を抑えることが出来る。また、システム内では複数の共有メモリおよび複数のキャッシュメモリを保持しているので、システム全体での共有メモリおよびキャッシュメモリへのアクセス性能を向上させる
10 ことができる。

図面の簡単な説明

- 第1図は、本発明を適用した計算機システムの構成例を示す図である。
- 15 第2図は、記憶装置の構成を示す図である。
- 第3図は、キャッシュメモリの管理情報および管理方法を示す図である。
- 第4図は、リードヒット処理の処理概要を示す図である。
- 第5図は、中央処理装置からリード処理要求を受領したホスト側プロ
20 セッサの処理のフローチャートである。
- 第6図は、記憶装置側プロセッサから再接続要求を受領したホスト側プロセッサの処理のフローチャートである。
- 第7図は、リード処理における記憶装置側プロセッサの処理のフローチャートである。
- 25 第8図は、ライト処理におけるホスト側プロセッサの処理のフローチャートである。

第 9 図は、ライト処理における記憶装置側プロセッサの処理のフローチャートである。

発明を実施するための最良の形態

5 以下、図面を用いて本発明を実施するための最良の形態を説明する。

第 1 図に示すように、計算機システムは、中央処理装置 100、制御装置 180、記憶装置 170 から構成される。制御装置 180 は、キャッシュメモリ 140 と中央処理装置 100 間のデータ転送を制御する一つ以上のホスト側プロセッサ 111 と、記憶装置 170 とキャッシュメモリ 140 間のデータ転送を制御する一つ以上の記憶装置側プロセッサ 152 から構成されている。

また、同一の記憶装置 170 にアクセス可能な全ての記憶装置側プロセッサ 152 を同一のプロセッサグループ 190 としており、各グループ毎に、一つの共有メモリ 160 とキャッシュメモリ 140 を設ける。

15 各グループ 190 内の記憶装置側プロセッサ 152 は、共有メモリアクセスパス 165 を介して共有メモリ 160 へアクセス可能である。各記憶装置側プロセッサ 152 は、記憶装置アクセスパス 175 を介して記憶装置 170 へアクセスする。記憶装置アクセスパス 175 は、一本のパス上に複数の記憶装置 170 が芋蔓状に接続されているものとする。

20 また、説明を簡単にするために、各記憶装置側プロセッサ 152 からは、一本の記憶装置アクセスパス 175 しか表記していないが、一つの記憶装置側プロセッサ 152 には、複数の記憶装置アクセスパス 175 が接続されていてもかまわない。

次に、第 2 図を用いて記憶装置 170 について説明する。本発明では

25 記憶装置 170 は、回転する記憶媒体を備えた磁気記憶装置であるとする。記憶装置 170 は、複数の記憶媒体 400、記憶媒体 400 毎に存

在し、記憶媒体 4 0 0 と記憶装置インタフェース 4 4 0 間のデータ転送を行う複数のヘッド 4 2 0、ヘッド 4 2 0 と記憶装置外部とのデータ転送を実行する記憶装置インタフェース 4 4 0 から構成される。

ヘッド 4 2 0 は上から昇順に番号がつけられている。複数の記憶媒体
5 4 0 0 から同時にデータをリード、ライトすることはできない。つまり、アクティブになるヘッド 4 2 0 は高々一つである。目的のデータをアクセスするために記憶媒体上をヘッド 4 2 0 を移動させることをシークと呼び、シークが完了した後に、目的データがヘッド 4 2 0 の下を通過するまで待つことをサーチと呼ぶ。ヘッド 4 2 0 が一回転する間にアク
10 セス可能な記憶媒体 4 0 0 上の領域をトラック 4 1 0 と呼ぶ。また、記憶媒体 4 0 0 が一回転する間に、全ヘッド 4 2 0 下を通過する領域（トラック 4 1 0 の集まり）をシリンダ 4 3 0 と呼び、記憶媒体 4 0 0 の外側から内側に向かって、昇順に番号が付けられている。シリンダ番号とヘッド番号の組みによって、トラック 4 1 0 を特定する。さらに、記憶
15 媒体 4 0 0 のデータ格納領域は 5 1 2 バイトの小領域に分かれていて、これをセクタと呼ぶ。セクタには、トラック 4 1 0 内で一意に番号が付けられていて、トラック 4 1 0 の先頭から昇順に、トラック 4 1 0 内で一意に番号が付けられている。

次に、第 3 図を用いてキャッシュメモリ 1 4 0 の管理方法を説明する
20 。キャッシュメモリ 1 4 0 は、各グループ 1 9 0 毎に分割されているが、論理的には連続したアドレス空間として見えるものとする。つまり、あるキャッシュアドレスが定まれば、それはどのグループ 1 9 0 に属するどの領域かを特定できるものとする。この実現のためには、例えばキャッシュアドレス上にグループ 1 9 0 の番号を埋め込んでおき、データ
25 転送を実行するハードウェアでそれを認識すれば良い。

キャッシュメモリ 1 4 0 は、セグメント 5 3 0 と呼ばれる例えば 1 6

K B（システムにより固定の大きさ）の領域に分割されていて、1 セグメントにつき、1 個のセグメント管理テーブル 5 2 0 が共用メモリ 1 6 0 上に存在する。セグメント管理テーブル 5 2 0 は、当該セグメント 5 3 0 のキャッシュアドレス 5 2 2、ダーティデータを示すビットマップ 5 2 3、クリーンデータを示すビットマップ 5 2 4、他セグメント管理テーブルへのポインタ 5 2 1 が格納されている。ここで、ダーティデータとは、キャッシュ 1 4 0 上は更新されているが、まだ記憶装置 1 7 0 に未反映のライトデータ、クリーンデータとは、キャッシュ 1 4 0 上に存在する記憶装置 1 7 0 の内容と一致したデータを表す。さらに、セグメント 5 3 0 を例えば 5 1 2 バイト毎に区切り、この領域に 1 ビットを対応させたビットマップを作成しておく。そして、もし対応する領域にダーティデータあるいはクリーンデータが存在するなら、ダーティビットマップ 5 2 3 あるいはクリーンビットマップ 5 2 4 の対応するビットを 1 にする。以上のようにより細かくデータ領域を管理するために、ダーティビットマップ 5 2 3 およびクリーンビットマップ 5 2 4 を使用する。

記憶装置 1 7 0 上の連続した領域、あるいは近傍の領域に対応してキャッシュ領域を割り当てる場合、複数のセグメント 5 3 0 をまとめて管理したほうが都合がよい。本発明では、1 トラック 4 1 0 分に対応するキャッシュ領域を、スロット管理テーブル 5 1 0 によって管理する。ここで、スロット管理テーブル 5 1 0 は、記憶装置番号 5 1 2、トラック番号 5 1 3、最初のセグメント管理ブロックへのポインタ 5 1 5、ロック情報 5 1 4 を格納している。記憶装置番号 5 1 2 およびトラック番号 5 1 3 によって、どの記憶装置のどのトラック 4 1 0 のデータを格納しているのかを特定でき、さらにポインタ 5 1 5 をたどり、セグメント管理ブロック 5 2 0 を参照することによって、そのトラック 4 1 0 のどの

セクタを格納しているのか、またそのセクタがダーティかどうかを特定できる。また、ロック情報 5 1 4 は、当該スロットを排他的に処理する場合に用いる。

ここで、ある記憶装置 1 7 0 のあるトラック 4 1 0 上のあるデータに
5 アクセス要求があった場合に、どのようにキャッシュ管理が行われるか
一例を示す。まず、アクセス要求として、リード・ライト種別、論理記憶装置番号、トラック番号、目的データの先頭のセクタ、長さ（セクタ数）が与えられる。まず、記憶装置番号とトラック番号を元に、目的のスロット管理ブロック 5 1 0 が存在するかどうかを調べる。この手法の一つとして、例えばハッシュがあげられる。すなわち、記憶装置番号と
10 トラック番号をハッシュ関数に与えると、対応するハッシュテーブル 5 0 0 のエントリを出力し、このエントリにスロット管理テーブル 5 1 0 のアドレスが格納されている。もしこれがヌル（アドレスなし）であれば、ミスである。ヌルでなければこのポインタをたどって、記憶装置番号
15 5 1 2 とトラック番号 5 1 3 が探しているものと一致しているかどうか調べる。一致してなければ、スロット管理ブロック 5 1 0 へのポインタをたどり、記憶装置番号 5 1 2 とトラック番号 5 1 3 の比較を繰り返す。このようにして、目的とするスロット管理ブロック 5 1 0 が存在するかどうかを探す。

20 もしスロット管理ブロック 5 1 0 が共有メモリ 1 6 0 上に存在すれば、続いてアクセス対象となっているセクタ範囲が存在するかどうかを調べる。そのためには、まず目的のデータの先頭セクタ番号とセクタ数から、先頭セグメント番号と先頭のビット位置、および最終セグメントと最終のビット位置を計算する。例えば、6 0 番のセクタを先頭に、1 2
25 個のセクタを読み出す場合では、第 2 セグメントのビット 2 9 から、第 3 セグメントのビット 9 が算出される。これを元にして、このビット範

囲のデータが存在するかどうかを、クリーンビットマップ 5 2 4 あるいはダーティビットマップ 5 2 3 を調べて判定する。もし目的の範囲全てに渡ってデータが存在すればヒット、もし一部分でも存在しなければミスということになる。ヒットの場合には、そのままアクセスを続行すれば良い。ミスの場合は、スロット管理テーブル 5 1 0 が存在しない場合と、スロット管理テーブル 5 1 0 は存在するが、セグメント管理テーブル 5 2 0 の一部あるいは全部が存在しない場合、セグメント管理テーブル 5 2 0 も全て存在するが、データがキャッシュ 1 4 0 上に存在しない場合がある。データがキャッシュ 1 4 0 上に存在しない場合には、データを記憶装置 1 7 0 からキャッシュ 1 4 0 へ読み込んでくれば良い。セグメント管理テーブル 5 2 0 が存在しない場合には、新しいセグメント管理テーブル 5 2 0 を必要数割り当て、スロット管理テーブル 5 1 0 に接続する。この際、ダーティビットマップ 5 2 3 とクリーンビットマップ 5 2 4 は初期化しておく。スロット管理テーブル 5 1 0 が存在しない場合には、スロット管理テーブル 5 1 0 一つと、セグメント管理テーブル 5 2 0 を必要数分割り当て、対応するハッシュテーブル 5 0 0 のエントリにスロット管理テーブル 5 1 0 のアドレスを格納する。この際、論理記憶装置番号 5 1 2 とトラック番号 5 1 3 を更新しておく。

目的のスロットがヒットであると判明した時点、あるいはミスであることが判明し、スロット管理テーブル 5 1 0 を新規に割り当てた時点で、スロットロックを確保する。もし、すでにスロットロック済みであった場合、当該スロットのロックが解放されるまで待つことになる。キャッシュ操作が完了し、もはや当該スロットが不要となると、スロットロックを解放する。この場合にも、ヒットミス判定時と同様に、論理記憶装置番号、トラック番号から目的のスロットをサーチする。あるいはヒットミス判定時に当該スロットのアドレスを記憶しておき、解放時はそ

のアドレスからスロット管理テーブル510を求める方法でもよい。目的とするスロット管理テーブル510を求めると、スロットロックを解除して、スロット解放処理が完了する。

以下、中央処理装置100からのアクセス要求に対する処理方法について説明する。中央処理装置100と制御装置180間のデータアクセスプロトコルは、汎用機で使用されるCKDプロトコル、ワークステーション等で使用されるSCSIプロトコルなどがあるが、本実施例ではCKDプロトコルを前提として説明する。

CKDプロトコルで使用されるコマンドのうち、本発明に関連するものは、アクセス有効範囲やキャッシュアクセスモードなどを示すDX (Define Extent)コマンド、アクセス位置やアクセスするレコード数を示すLOC (Locate)コマンド、リードを行うRDD (Read Data)コマンド、ライトを行うWRD (Write Data)コマンドなどがある。あるレコードをアクセスする場合には、DX, LOC, RDDなど、複数のコマンドが連続して実行される。これをコマンドチェーンと呼ぶ。また、DX, LOC, RDD...とコマンド発行することにより、複数の連続したレコードをアクセスすることもできる。DXコマンドでは、レコードを逐次的に読み出すかどうかを指定するモードがある。LOCコマンドによって転送される位置づけ情報は、シリンダ番号、ヘッド番号、セクタ番号、レコード番号が指定される。また、記憶装置番号は中央処理装置がコマンドを発行するために、制御装置との接続を確立する際に指定される。

まず、第4図を用いて処理全体の概要を説明する。

第4図は、リードヒット処理の概略図を示す。ホスト側プロセッサ111が、中央処理装置100からのリード要求を受領すると、リード対象となった記憶装置170にアクセス可能な記憶装置側プロセッサ15

2を以下の方法で選択し、リードメッセージを送信する。

各プロセッサは、そのローカルメモリ上に記憶装置アクセス表600を保持している。各エントリは、各該記憶装置170にアクセス可能な記憶装置側プロセッサ番号と状態の組から構成され、この例では各記憶装置170に高々2つの記憶装置側プロセッサ152がアクセス可能な構成となっている。状態は正常あるいは閉塞のいずれかであり、閉塞状態の時は、当該プロセッサ152から記憶装置170へのアクセスはできない。この表は、システム構築時に共有メモリ160に作成されて、各プロセッサはそのコピーをローカルメモリ上に保持する。障害により記憶装置側プロセッサ152と記憶装置間アクセスパス175が閉塞した場合や、保守により一時的にアクセス不能にする場合は、本記憶装置アクセス表600の対応するエントリの状態を閉塞とする。

ホスト側プロセッサ111は、記憶装置アクセス表600の、アクセス対象である記憶装置番号に対応するエントリを見て、当該記憶装置170にアクセス可能な記憶装置側プロセッサ152を求め、任意の記憶装置側プロセッサに対して、処理要求のメッセージを発行する。メッセージは、アクセス種別、記憶装置番号、シリンダ番号、ヘッド番号、セクタ番号、セクタ数から成る。ここで、アクセス種別からセクタ番号までは中央処理装置から指定される。セクタ数は、LOCコマンドで指定されたレコード数とレコード長から算出する。例えば、1セクタ512バイトで、4キロバイトのレコードを3個読み出す場合には、 $4096 \times 3 \div 512 = 24$ セクタとなる。

ホスト側プロセッサ111からのリード要求に対して、記憶装置側プロセッサ152は共有メモリをアクセスしてヒットミス判定を実行し、当該データを格納しているキャッシュ領域のアドレス、およびキャッシュ140上の有効データ長をホスト側プロセッサ111へ返す。これを

受けて、ホスト側プロセッサ 1 1 1 は当該アドレスからデータを読み出して、中央処理装置 1 0 0 へデータ転送を行う。以上の処理が完了すると、中央処理装置 1 0 0 へ正常終了を報告するとともに、記憶装置側プロセッサ 1 5 2 へアクセス完了を報告し、処理を終了する。これを受けて記憶装置側プロセッサ 1 5 2 は、確保していたキャッシュ領域を解放する。

以上、中央処理装置 1 0 0 からのリード要求に対する処理方式の概要を説明した。以下、リードおよびライト処理の詳細を処理フロー図を用いて説明する。

10 まず、リード処理方式について、第 5 図から第 7 図を用いて説明する。

第 5 図はホスト側プロセッサ 1 1 1 の処理を示す。ステップ 7 0 0 で中央処理装置 1 0 0 からリード処理要求を受領すると、ステップ 7 2 0 において、上述した方法で、記憶装置側プロセッサ 1 5 2 を選択し、リード処理要求を発行する。ステップ 7 3 0 で、記憶装置側プロセッサ 1 5 2 からの応答を待つ。

第 7 図のステップ 9 0 0 において、記憶装置側プロセッサ 1 5 2 がホスト側プロセッサ 1 1 1 からの処理要求を受領すると、ステップ 9 1 0 でヒットミス判定を行う。その結果、ヒットであることがわかると、ステップ 9 2 0 でメッセージを発行したホスト側プロセッサ 1 1 1 に対して、以下の報告する。

- (1) ヒットミス判定結果：ヒット。
- (2) セグメントアドレスリスト：データを格納しているセグメントアドレスのリスト。
- 25 (3) セグメント内オフセット：データを格納している先頭セクタの先頭セグメント内オフセット。

ここで、セグメントアドレスはセグメント管理テーブル520に格納されている。また、セグメント内オフセットは、セグメント管理テーブル520のダーティビットマップ523あるいはクリーンビットマップ524における、目的データを格納した先頭セクタに対応するビット位置を返す。

5
10
15
20
25
ホスト側プロセッサ111がヒットであることを認識すると、前記セグメントアドレスリストおよびセグメント内オフセットから、目的のデータを格納したキャッシュアドレスがわかるので、ステップ740で中央処理装置100に対してデータを転送する。その後、ステップ750で記憶装置側プロセッサ152に対して、アクセス完了を報告し、処理を完了する。この時送信されるメッセージは、記憶装置番号、シリンダ番号、ヘッド番号、ダーティ有無情報、先頭ダーティセグメント、セグメント内オフセット、ダーティセクタ数から構成される。ここで、ダーティ有無情報はダーティなしが格納され、先頭ダーティセグメント、セグメント内オフセット、ダーティセクタ数は無効値が格納されている。第7図のステップ930で、記憶装置側プロセッサ152が、ホスト側プロセッサ111から、アクセス完了報告を受領すると、現在確保中のスロットのロックを解放して、処理を終了する。

もし、キャッシュミスしていた場合には、記憶装置側プロセッサ152は、第7図のステップ910のヒットミス判定処理において、新規にスロット管理テーブル510ならびにセグメント管理テーブル520を割り当てる。続いてステップ950でホスト側プロセッサ111に対してヒットミス判定結果（ミス）を報告した後、ステップ960で、データを格納した記憶装置170に対してリード要求を発行し、リード完了まで待つ。

ホスト側プロセッサ111は、第5図のステップ735において、リ

ードミスであることを認識すると、いったん処理を中断し、中央処理装置 100 との接続を切り離す（ステップ 770）。

記憶装置 170 からキャッシュメモリ 140 へのデータ転送が完了すると、記憶装置側プロセッサ 152 は、第 7 図のステップ 970 において、当該記憶装置 170 からの読み出しを待っているホスト側プロセッサ 152 に対して、再接続要求を発行する。

第 6 図のステップ 800 において、ホスト側プロセッサ 111 が記憶装置側プロセッサ 152 からの再接続要求を検出すると、ステップ 810 において、ホスト側プロセッサ 111 と中央処理装置 100 間で再接続処理を実行する。これを完了すると、ステップ 830 で、記憶装置側プロセッサ 152 に対して再接続完了を報告する。

これを受けて、記憶装置側プロセッサ 152 は、第 7 図のステップ 990 で、要求データを格納したセグメントアドレスリスト、セグメント内オフセットを送信する。以降の処理はリードヒットの処理と同様なので省略する。

もしヒットミス判定の結果がライトミスだった場合は、ホスト側プロセッサ 111 からの要求が、ミス時リード要か不要かによって、対応が異なる。もし、ミス時リード要であった場合には、前記リードミス時と同様に、記憶装置 170 からキャッシュ 140 へデータを格納する。ホスト側プロセッサ 111 ではこの間、中央処理装置 100 との接続を切り離しており、再開処理にてライト処理を実行する。ここでのライト処理はヒット時のライト処理と同様である。ミス時リード不要であれば、記憶装置側プロセッサ 152 およびホスト側プロセッサ 111 共、ライトヒット処理と同様の処理を行う。

次に、ライト処理方法について、第 8 図、第 9 図を用いて説明する。

第 8 図のステップ 1000 において、ホスト側プロセッサ 111 が中

中央処理装置 1 0 0 からライト処理要求を受領すると、ステップ 1 0 2 0 で、第 4 図で説明した方法で記憶装置側プロセッサ 1 5 2 を選択して、ライト処理要求を発行し、ステップ 1 0 3 0 で記憶装置側プロセッサ 1 5 2 からの応答を待つ。

- 5 記憶装置側プロセッサ 1 5 2 では、第 9 図のステップ 1 1 1 0 でヒットミス判定を実行する。ここで、もしミスであれば、ヒットミス判定処理で、スロット管理テーブル 5 1 0 およびセグメント管理テーブル 5 2 0 を必要分割り当てる。続いてステップ 1 1 2 0 で、セグメントアドレスリスト、セグメント内オフセットをホスト側プロセッサへ転送した後、
- 10 ステップ 1 1 3 0 でホスト側プロセッサ 1 1 1 からのスロット解放要求を待つ。

- ホスト側プロセッサ 1 1 1 では、第 8 図のステップ 1 0 4 0 で、記憶装置側プロセッサ 1 5 2 から、ヒットミス判定結果、およびキャッシュアドレスを受け取ると、当該キャッシュ領域に、中央処理装置 1 0 0 から
- 15 転送されたデータを格納する。データ転送が完了すると、ステップ 1 0 5 0 において、ホスト側プロセッサ 1 1 1 は記憶装置側プロセッサ 1 5 2 に対して、アクセス完了を送信する。この時送信されるメッセージは、記憶装置番号、シリンダ番号、ヘッド番号、ダーティ有無情報、先頭ダーティセグメント、セグメント内オフセット、ダーティセクタ長から
- 20 構成される。ここで、ダーティ有無情報はダーティありが格納され、先頭ダーティセグメントおよびセグメント内オフセットには、先頭のダーティセクタを保持したセグメントアドレスおよびそのセグメント内のセクタオフセット、ダーティセグメント数には中央処理装置 1 0 0 からライトされたセグメントの個数が格納される。

- 25 アクセス完了報告を受領した記憶装置側プロセッサ 1 5 2 は、第 9 図のステップ 1 1 4 0 で、前述の方法で目的とするスロット管理テーブル

5 1 0 を取得すると、先頭ダーティセグメントに対応するセグメント管理テーブル 5 2 0 のダーティビットマップ 5 2 3 の、セグメント内オフセットに対応するビットから、ダーティセグメント数分のビットを 1 にする（ステップ 1 1 4 0）。この際、ダーティセクタ数によっては複数
5 セグメントにわたってダーティビットマップ 5 2 3 を 1 にする場合もありうる。この場合には次セグメントポインタをたどって、ダーティビットマップを順次 1 にしていく。この処理を完了すると、当該スロットロックを解除し、ホスト側プロセッサ 1 1 1 に完了報告を行った後、記憶装置 1 7 0 に対してライト要求を発行し、データを書き込む。

10

産業上の利用可能性

以上のように、本発明にかかる記憶制御装置およびその制御方法は、複数の記憶装置側プロセッサと共有メモリとキャッシュメモリによりグループを構成し、グループ内だけで共有メモリへのアクセスを許可することによって、制御を単純化して共有メモリアクセスネックを低減し、
15 かつ共有メモリおよびキャッシュメモリを各グループ毎に分散して持つことによって、記憶制御装置全体のスループットを向上することができる記憶制御装置およびその制御方法を構築するのに適している。

請 求 の 範 囲

1. 中央処理装置と複数の記憶装置とに接続され、前記中央処理装置と記憶装置との間でのデータ転送を実行する記憶制御装置の制御方法であって、

- 5 前記複数の記憶装置をグループ分けし、当該グループ対応にキャッシュメモリ及び当該キャッシュメモリの管理情報を格納する共有メモリを設け、

各グループ対応に設けられた記憶装置側プロセッサにより、当該グループに属する記憶装置と当該グループに割り当てられた前記キャッシュとの間のデータ転送を実行し、

10 ホスト側プロセッサにより、前記中央処理装置がアクセスを要求した記憶装置が属するグループを識別し、ホスト側プロセッサが前記中央処理装置と当該グループに割り当てられたキャッシュとの間のデータ転送を実行することを特徴とする記憶制御装置の制御方法。

- 15 2. 前記記憶装置側プロセッサは、当該記憶装置側プロセッサが属さないグループ内の記憶装置のデータのキャッシュアドレスを、当該記憶装置が属するグループの記憶装置側プロセッサから通知を受けることを特徴とする請求の範囲第1項記載の記憶制御装置の制御方法。

3. 前記ホスト側プロセッサは、前記中央処理装置からのリード要求に対して、前記記憶装置側プロセッサへ、記憶装置アドレス、データ位置、

20 リード範囲とともにリード要求を発行し、
前記記憶装置側プロセッサは、共有メモリをアクセスして、目的データがキャッシュ上に存在するかどうか判定し、

- 存在する場合は、前記記憶装置側プロセッサは、前記目的データのキャッシュアドレスと有効データ長を前記ホスト側プロセッサに対して通知し、

前記ホスト側プロセッサは、前記キャッシュアドレスからデータを読み出して、前記中央処理装置にデータ転送を行うことを特徴とする請求の範囲第1項記載の記憶制御装置の制御方法。

4. 前記目的データがキャッシュ上に存在しない場合、記憶装置側プロセッサはリードミス을ホスト側プロセッサへ通知するとともに、記憶装置に対してリード要求を発行し、

前記ホスト側プロセッサは前記記憶装置側プロセッサおよび前記中央処理装置との接続を切り離し、記憶装置からキャッシュへデータを転送し、前記記憶装置側プロセッサから再接続要求が発行されたことに応じて、前記中央処理装置と再接続処理を実行し、前記記憶装置側プロセッサに対して再接続完了を報告し、

前記記憶装置側プロセッサは、前記データが格納されたキャッシュアドレスおよび有効データ長を前記ホスト側プロセッサに通知し、

- 前記ホスト側プロセッサは、前記キャッシュアドレスからデータを読み出して、前記中央処理装置にデータ転送を行うことを特徴とする請求の範囲第1項記載の記憶制御装置の制御方法。

5. 前記中央処理装置が前記制御装置に対してライト要求を発行した場合、前記ライト要求を受領した前記ホスト側プロセッサは、前記記憶装置側プロセッサに対して、記憶装置アドレス、データ位置、ライト範囲とともにライト要求を発行し、

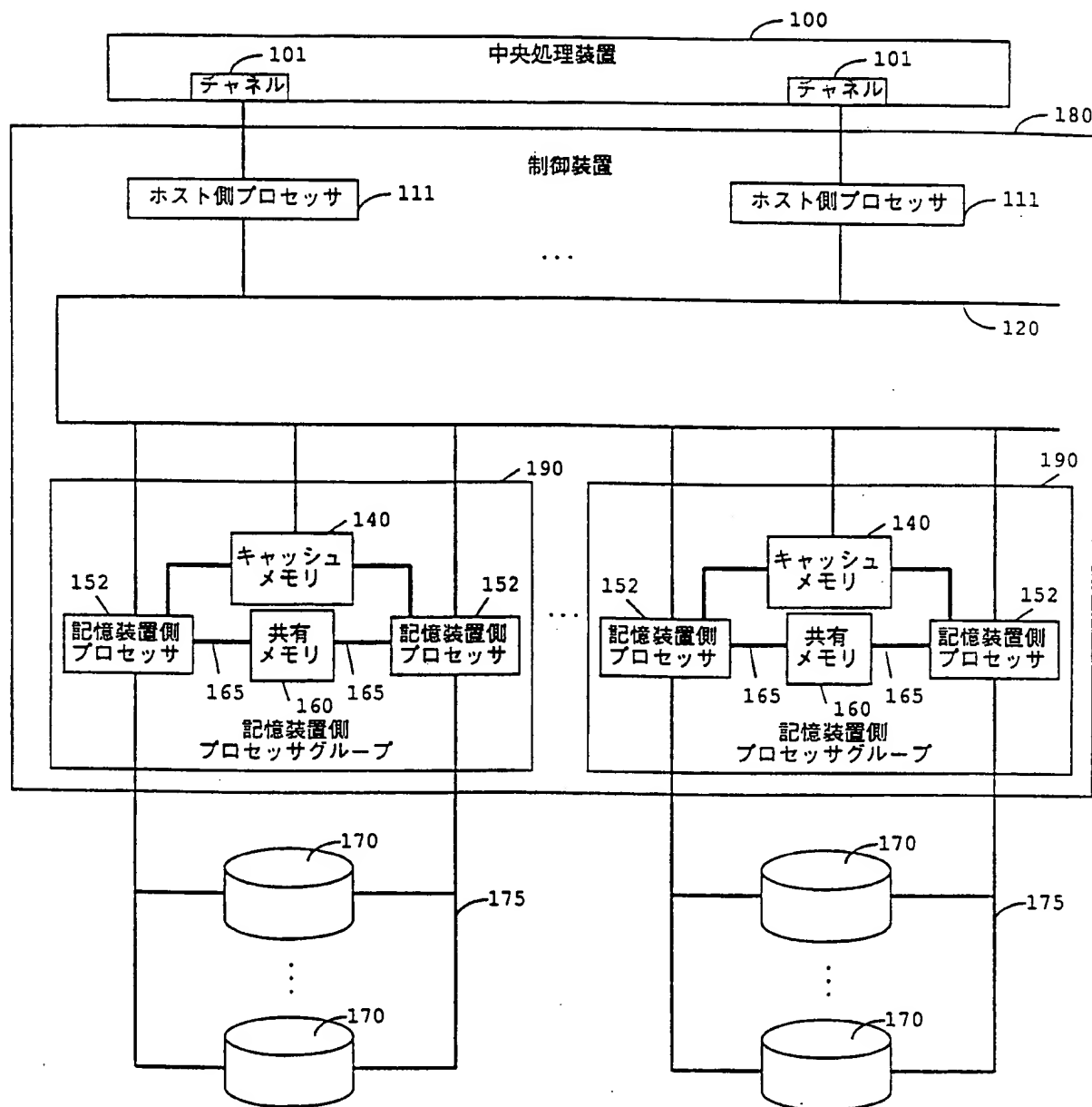
前記ライト要求を受領した前記記憶装置側プロセッサは、ライトデータを格納するためのキャッシュ領域を確保し、データを格納すべきキャッシュアドレスを前記ホスト側プロセッサへ通知し、

- 前記ホスト側プロセッサは、前記キャッシュアドレスに前記中央処理装置から転送されたデータを格納し、前記記憶装置側プロセッサに対して、ライトデータを格納しているキャッシュアドレスならびにデータ長

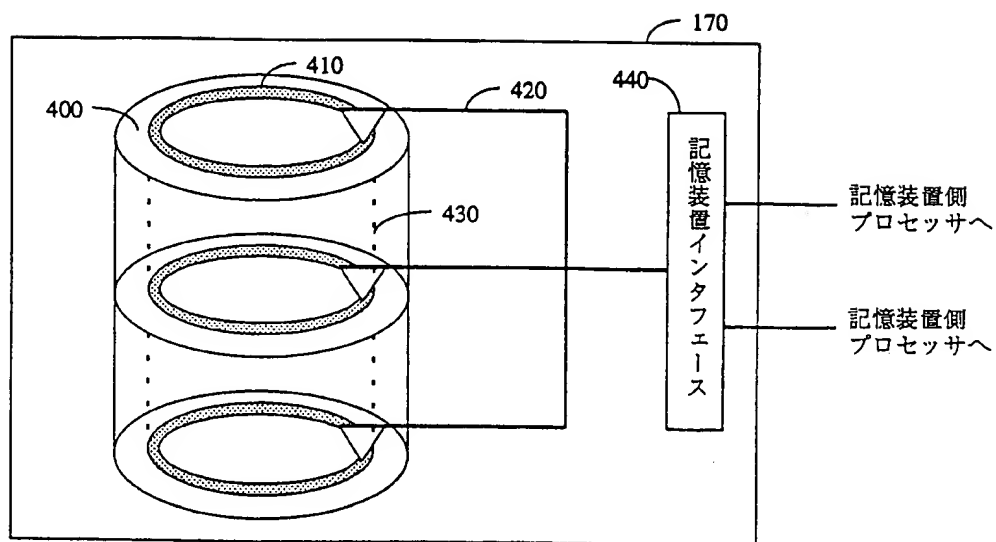
を送信し、

前記記憶装置側プロセッサは、前記キャッシュに格納されたライトデータを前記記憶装置に書き込むことを特徴とする請求の範囲第1項記載の記憶制御装置の制御方法。

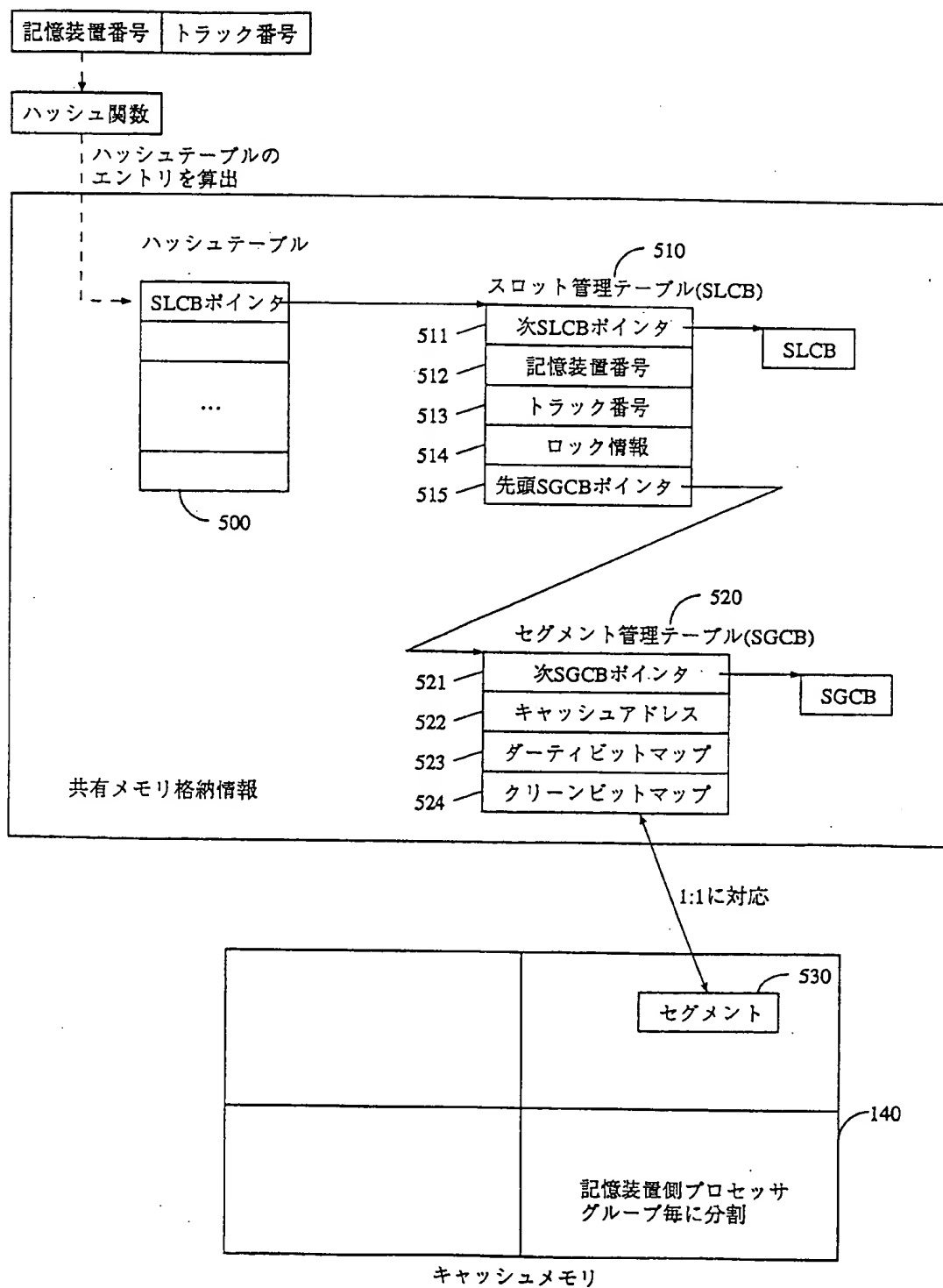
第 1 図



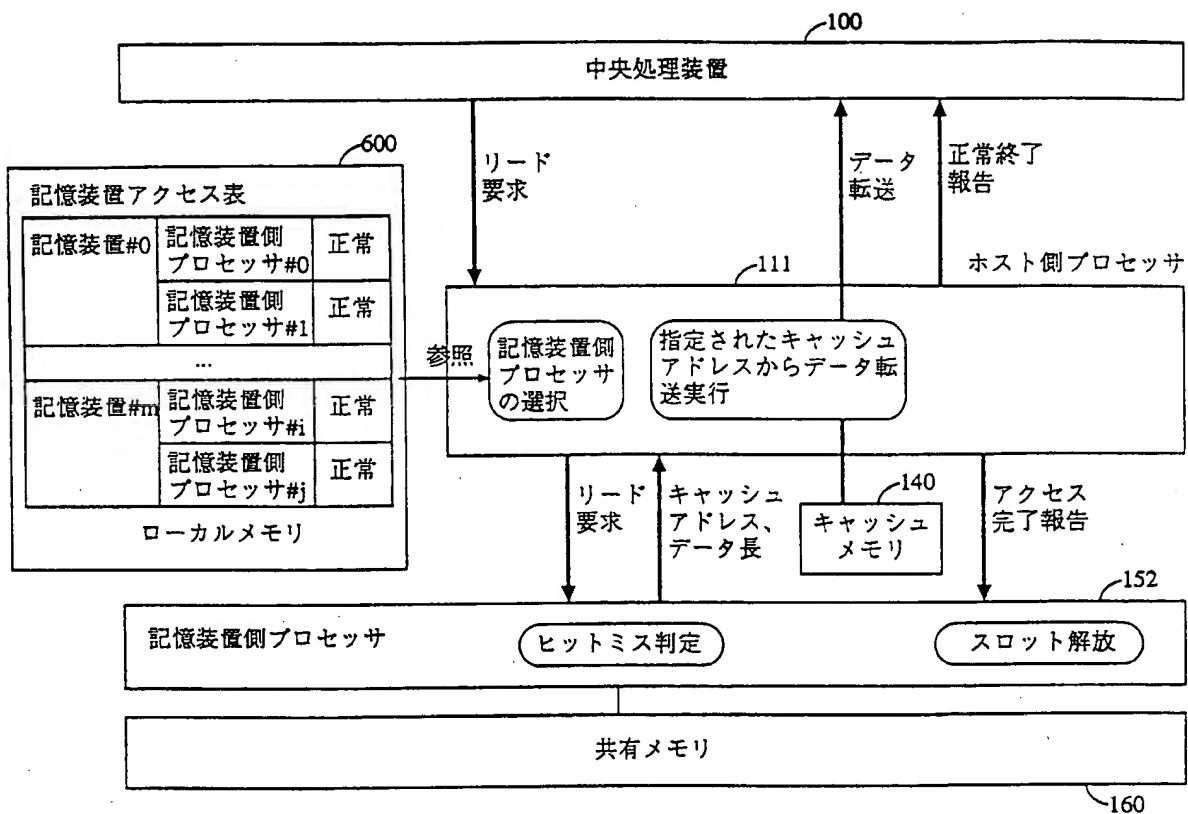
第 2 図



第 3 図

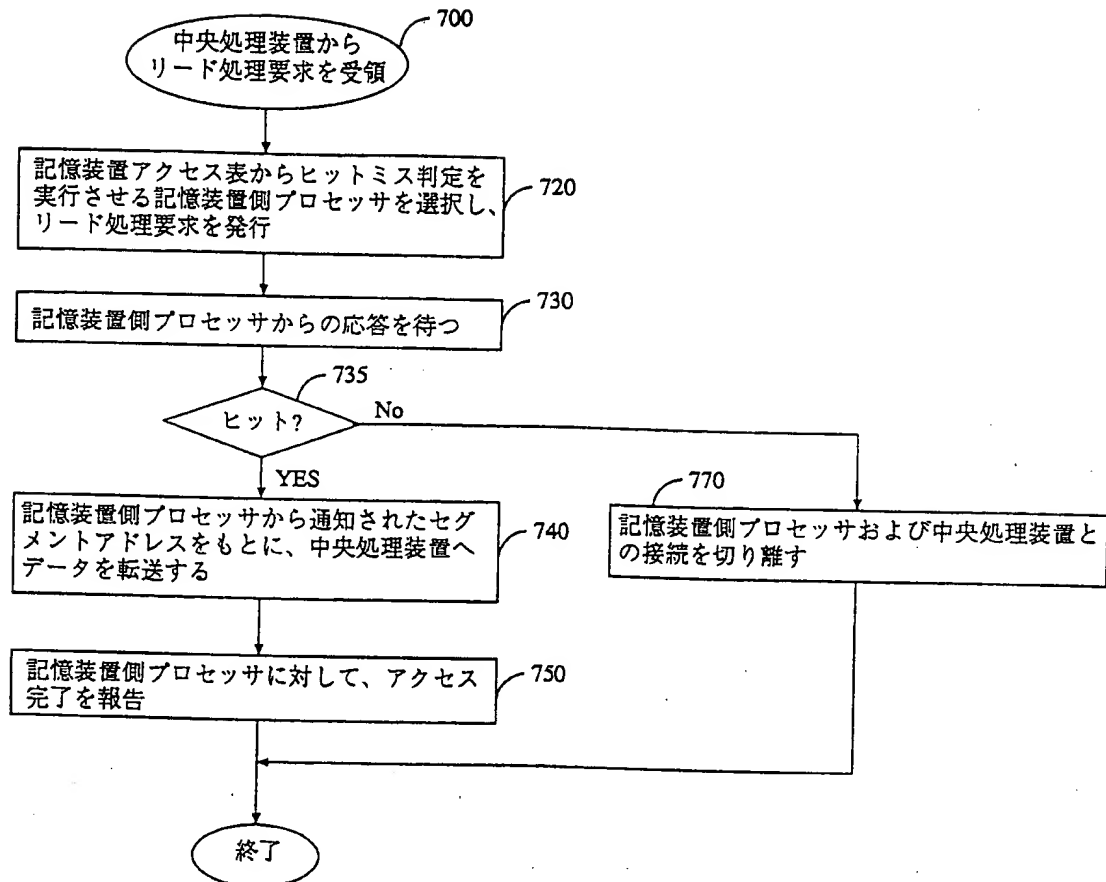


第 4 図



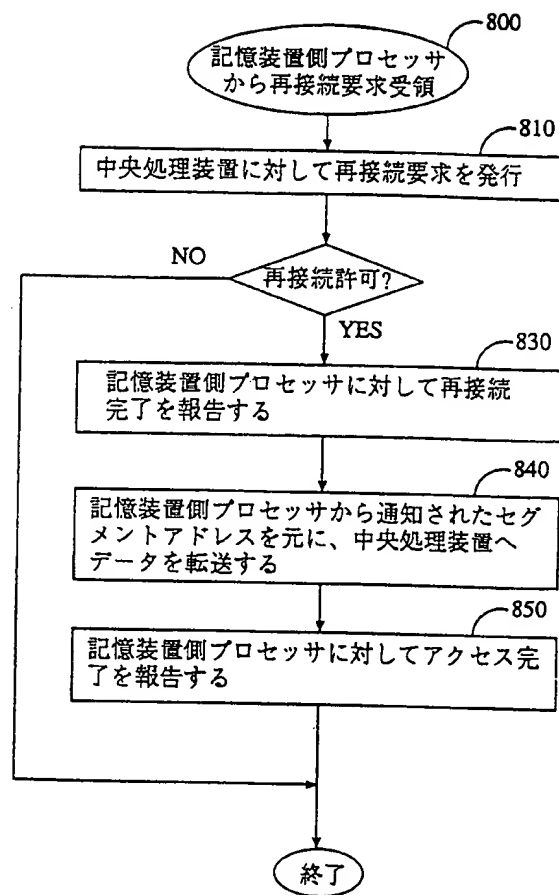
5 / 9

第 5 図

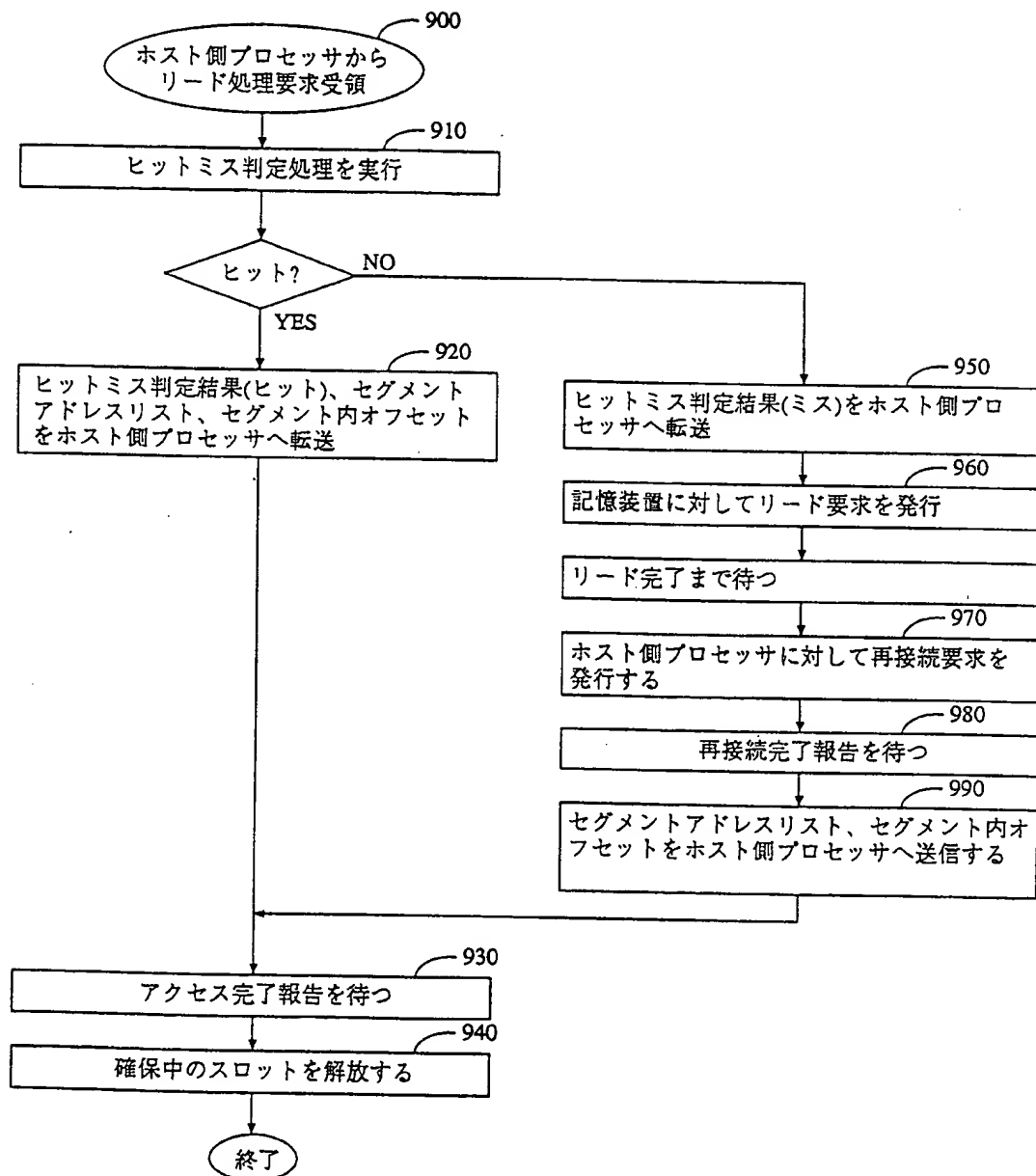


6 / 9

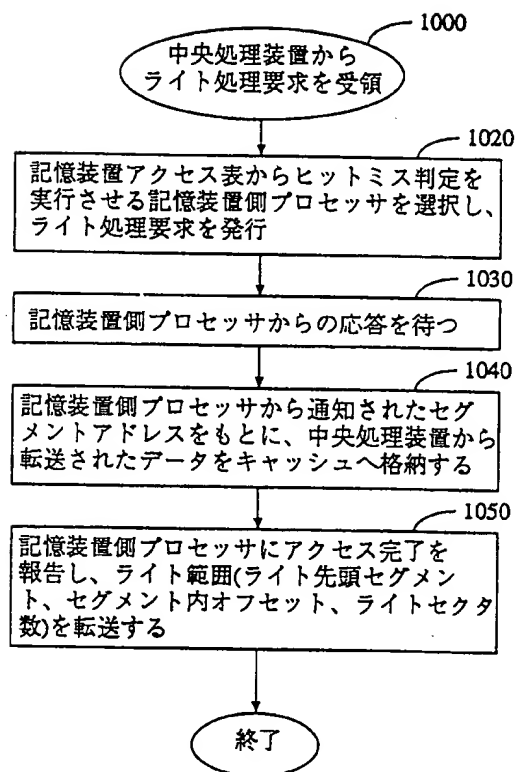
第 6 図



第 7 図

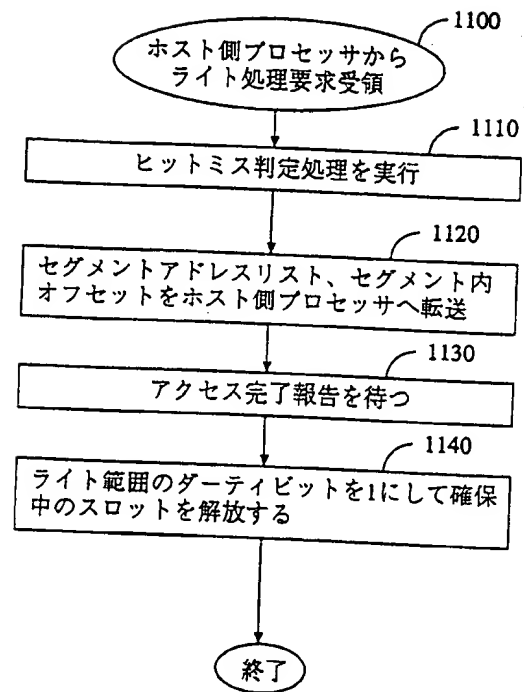


第 8 図



9 / 9

第 9 図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP97/02135

A. CLASSIFICATION OF SUBJECT MATTER

Int. C1⁶ G06F3/06

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. C1⁶ G06F3/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho

1926 - 1997

Kokai Jitsuyo Shinan Koho

1971 - 1997

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|--|-----------------------|
| X Y | JP, 59-117653, A (NEC Corp.), July 7, 1984 (07. 07. 84) (Family: none) | 1 2 - 5 |
| A | JP, 2-68639, A (Hitachi, Ltd.), March 8, 1990 (08. 03. 90) (Family: none) | 3, 4 |
| A | JP, 60-73758, A (Hitachi, Ltd.), April 25, 1985 (25. 04. 85) (Family: none) | 5 |

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

July 11, 1997 (11. 07. 97)

Date of mailing of the international search report

July 29, 1997 (29. 07. 97)

Name and mailing address of the ISA/

Japanese Patent Office

Facsimile No.

Authorized officer

Telephone No.

| | | |
|---|---|--|
| A. 発明の属する分野の分類 (国際特許分類 (IPC)) | | |
| Int. Cl ⁸ G06F 3/06 | | |
| B. 調査を行った分野 | | |
| 調査を行った最小限資料 (国際特許分類 (IPC)) | | |
| Int. Cl ⁸ G06F 3/06 | | |
| 最小限資料以外の資料で調査を行った分野に含まれるもの | | |
| 日本国公開実用新案公報 1971-1997年 日本国実用新案公報 1926-1997年 | | |
| 国際調査で使用した電子データベース (データベースの名称、調査に使用した用語) | | |
| C. 関連すると認められる文献 | | |
| 引用文献の カテゴリー* | 引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示 | 関連する 請求の範囲の番号 |
| X Y | JP, 59-117653, A (日本電気株式会社) 7. 7月. 1984 (07. 07. 84) (ファミリー無し) | 1 2~5 |
| A | JP, 2-68639, A (株式会社日立製作所) 8. 3月. 1990 (08. 03. 90) (ファミリー無し) | 3, 4 |
| A | JP, 60-73758, A (株式会社日立製作所) 25. 4月. 1985 (25. 04. 85) (ファミリー無し) | 5 |
| <input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。 | | |
| * 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 先行文献ではあるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献 | | |
| 国際調査を完了した日 11. 07. 97 | | 国際調査報告の発送日 29.07.97 |
| 国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100 東京都千代田区霞が関三丁目4番3号 | | 特許庁審査官 (権限のある職員) 井出 和水 印 電話番号 03-3581-1101 内線 3523 |

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.